Лекция 10. Задачи с ограничениями

x_* — решение

$$\min_{\substack{Ax=b\\ x\in Q}} f(x)$$

f — вып.

Q — мн-во простой структуры

$$x^{k+1} = \pi_Q(x^k - h\underbrace{\nabla f(x^k)}_{\text{дороже чем}})$$

Пример. $Q = \mathbb{R}^d_+$

$$x^{k+1} = \pi_Q(x^k - h\nabla f(x^k)) = [x^k - h\nabla f(x^k)]_+$$

$$f(x) = \frac{1}{2}\langle x, Cx\rangle - \langle d, x\rangle$$

$$\nabla f(x) = \underbrace{Cx - d}_{\text{дороже, чем } [\,..\,]_+}$$

Пример. $Ax = b \to g(x) \le 0$, g — вып.

$$\max_\lambda \langle \lambda, Ax-b\rangle$$

$$\downarrow$$

$$\max_{\lambda\ge 0} \langle \lambda, g(x)\rangle$$

1 стратегия (седловая)

$$\min_{\substack{Ax=b\\ x\in Q}} f(x) = \min_{x\in Q}\Big\{f(x) + \underbrace{\max_\lambda \langle\lambda, Ax-b\rangle}_{=\begin{cases}0, & Ax=b\\ +\infty, & Ax\ne b\end{cases}}\Big\} =$$

$$= \min_{x\in Q}\max_\lambda \underbrace{f(x) + \langle \lambda, Ax-b\rangle}_{L(x,\lambda)}$$

$L(x,\lambda)$: по x — вып., по λ — лин. $\Rightarrow$ вогн.

2 стратегия (двойств.) фон-Нейман, Сион–Какутани

$$\min_{x\in Q}\max_\lambda \underbrace{f(x) + \langle\lambda, Ax-b\rangle}_{L(x,\lambda)} = \max_\lambda \min_{x\in Q} L(x,\lambda)$$

$L(x,\lambda)$ – вып. по x, вог. по λ, Q – вып. и комп.

$$\max_{\lambda} \underbrace{\min_{x\in Q} \{ \langle \lambda, Ax-b\rangle + f(x) \}}_{\tilde{\varphi}(\lambda)}$$

$\varphi(\lambda) = -\tilde{\varphi}(\lambda)$

двойств. задача

$$\min_{\lambda} \varphi(\lambda)$$

$$\| $$

$$\max_{x\in Q} \{ \langle \lambda, b-Ax\rangle - f(x) \}$$

Пример. $f(x) = \sum_{k=1}^{d} f_k(x_k)$ → разделяется на 1D если $Q = \prod_{i=1}^{d} [a_i, b_i]$

$\nabla\varphi(\lambda) = b - Ax(\lambda)$, где

$x(\lambda)$ – реш. задачи $\max_{x\in Q} \{ \underbrace{\langle \lambda, b-Ax\rangle - f(x)}_{} \}$

$x(\lambda)$ – реш. этой задачи

$\varphi(\lambda) = \max_{x\in \mathbb{R}^d} L(x,\lambda)$

$L_x(x,\lambda) = 0$

$\Downarrow$

$x(\lambda)$

$$\nabla\varphi(\lambda) = \nabla_\lambda \Big(L(x(\lambda),\lambda) \Big) = \nabla x(\lambda)^T \cdot \underbrace{L_x(x(\lambda),\lambda)}_{0} + \underbrace{L_\lambda(x(\lambda),\lambda)}_{} =$$

$$= \frac{\partial L}{\partial \lambda}(x,\lambda) \Big|_{x=x(\lambda)}$$

$$\varphi(\lambda) = \max_{x\in Q} \{\langle\lambda, b - Ax\rangle - f(x)\}$$

$\tilde\lambda$ – приближ. реш. задачи $\min\limits_{\lambda} \varphi(\lambda)$

$x(\tilde\lambda)$ – какое качество у этого решения?

$$\langle\tilde\lambda, \underbrace{b - Ax(\tilde\lambda)}_{\nabla\varphi(\tilde\lambda)}\rangle - f(x(\lambda)) \geq \langle\tilde\lambda, \underbrace{b - Ax_*}_{=0}\rangle - f(x_*) = -f(x_*)$$

1) $f(x(\tilde\lambda)) - f(x_*) \le \langle\tilde\lambda, b - Ax(\tilde\lambda)\rangle = \langle\tilde\lambda, \nabla\varphi(\tilde\lambda)\rangle$

2) $\|Ax(\tilde\lambda) - b\|_2 = \|\nabla\varphi(\tilde\lambda)\|_2$

Чтобы
$$f(x(\tilde\lambda)) - f(x_*) \le \varepsilon$$
$$\|Ax(\tilde\lambda) - b\|_2 \le \tilde\varepsilon$$

достаточно
$$\langle\tilde\lambda, \nabla\varphi(\tilde\lambda)\rangle \le \varepsilon$$
$$\|\nabla\varphi(\tilde\lambda)\|_2 \le \tilde\varepsilon.$$

Другой пример

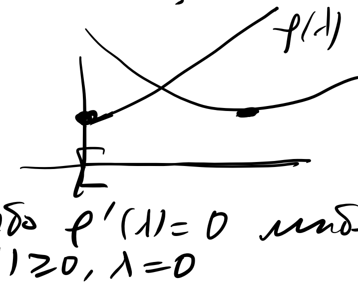
$\min\limits_{\substack{g(x)\le 0 \\ x\in Q}} f(x)$ → вып.

$g(x)\le 0$ → вып.

$$\varphi(\lambda) = \max_{x\in Q} \{-\langle\lambda, g(x)\rangle - f(x)\}$$

$$\nabla\varphi(\lambda) = -g(x(\lambda))$$

$$\min_{\lambda\ge 0} \varphi(\lambda)$$



скаляр. $\left[\begin{array}{l}\text{либо } \varphi'(\lambda) = 0 \text{ либо} \\ \varphi'(\lambda) \ge 0,\ \lambda = 0\end{array}\right.$

$$-\langle \tilde\lambda, g(x(\tilde\lambda))\rangle - f(x(\tilde\lambda)) \geq \underbrace{-\langle \tilde\lambda, g(x_*)\rangle}_{\geq 0} - f(x_*) \geq -f(x_*)$$

$\tilde\lambda \geq 0$, $g(x_*) \leq 0$

$$f(x(\tilde\lambda)) - f(x_*) \leq \langle \tilde\lambda, g(x(\tilde\lambda))\rangle =$$

$$= \langle \tilde\lambda, \nabla\varphi(\tilde\lambda)\rangle$$

1) $f(x(\tilde\lambda)) - f(x_*) \leq \langle \tilde\lambda, \nabla\varphi(\tilde\lambda)\rangle$

2) $\|[g(x(\tilde\lambda))]_+\|_2 = \|[-\nabla\varphi(\tilde\lambda)]_+\|_2$

$\tilde\lambda = \lambda_*$ – реш. двойств. задачи, то $x(\lambda_*) = x_*$.

Условие Слейтера

$\|\tilde\lambda\|_2$ – ?

<u>Ут.</u> $\exists\ \bar x \in Q$: $\gamma = \min\limits_i \{-g_i(\bar x)\} > 0$

$\forall\ i=\overline{1,m}\quad g_i(\bar x) \leq -\gamma$

$g(x) \leq 0 \Leftrightarrow g_i(x) \leq 0$, $i=\overline{1,m}$

вектор ф-ций

$$\varphi(0) \geq \varphi(\lambda_*) = \max_{x\in Q}\left\{-\sum_{i=1}^m \lambda_{*i}\, g_i(x) - f(x)\right\} \geq -\sum_{i=1}^m \lambda_{*i}\, g_i(\bar x) - f(\bar x)$$

по построению $\lambda_* = \arg\min\limits_{\lambda \geq 0} \varphi(\lambda)$

$$\varphi(0)+f(\bar{x}) \geq -\sum_{i=1}^{m} \lambda_{*i}\, g_i(\bar{x}) \geq \gamma \underbrace{\sum_{i=1}^{m} \lambda_{*i}}_{\geq 0} = \gamma \|\lambda_*\|_1$$

$$\|\lambda_*\|_1 \leq \frac{1}{\gamma}\left(f(\bar{x})+\varphi(0)\right) =$$

$$= \frac{1}{\gamma}\left(f(\bar{x}) - \min_{x\in Q} f(x)\right)$$

Другой пример.

$$\min_{\substack{Ax=b \\ x\in Q}} f(x)$$

$$\varphi(\lambda) = \max_{x\in Q} \{\langle \lambda, b-Ax\rangle - f(x)\} = \Phi(A^T\lambda)$$

$$-f(x_*) = \langle \lambda_*, \underbrace{b-Ax_*}_{=0}\rangle - f(x_*) =$$

$$= \max_{x\in Q} \{\langle \lambda_*, b-Ax\rangle - f(x)\} \geq$$

$$\geq \langle \lambda_*, b-Ax\rangle - f(x) =$$

$$= \langle \lambda_*, Ax_* - Ax\rangle - f(x) = \langle A^T\lambda_*, x_*-x\rangle - f(x)$$

$$f(x) \geq f(x_*) + \langle -A^T\lambda_*, x-x_*\rangle$$

Из выпуклости $f(x) \Rightarrow -A^T\lambda_* \in \partial f(x_*)$

$$-A^T\lambda_* = \nabla f(x_*)$$

$$\lambda_* \in (\operatorname{Ker} A^T)^\perp \qquad \|\nabla f(x_*)\|_2^2$$

$$\|-A^T\lambda_*\|_2^2 = \langle -A^T\lambda_*, -A^T\lambda_* \rangle =$$

$$= \langle \lambda_*, A^T A \lambda_* \rangle \geq \lambda_{min}^{+}(A^T A)\, \|\lambda_*\|_2^2$$

мин. полож. собств. знач. $A^T A$.

$$\boxed{\|\lambda_*\|_2^2 \leq \frac{\|\nabla f(x_*)\|_2^2}{\lambda_{min}^{+}(A^T A)}}$$

Возвращение к двойственной задаче

$$\min_{\lambda \in \mathbb{R}^m} \varphi(\lambda)$$

$$\varphi(\lambda) = \max_{x \in \mathbb{R}^d} \{ \langle \lambda, x \rangle - f(x) \}$$

$$f(x) = \frac{\mu}{2} x_1^2 + \frac{L}{2} x_2^2$$

$$\varphi(\lambda) = \frac{1}{2L}\lambda_1^2 + \frac{1}{2\mu}\lambda_2^2$$

$$\max_{x_i} \left\{ \lambda_i x_i - \frac{c x_i^2}{2} \right\} \qquad \lambda_i = c_i x_i \Rightarrow x_i = \frac{\lambda_i}{c_i}$$

$$\frac{\lambda_i^2}{c_i} - \frac{\lambda_i^2}{2c_i} = \frac{\lambda_i^2}{2c_i}$$

$f(x) \nearrow$ μ-сильно выпукл. в 2-норме, то

$\searrow$ L-липш. град.

$Q = \mathbb{R}^d$ важно для $\to$ $\frac{1}{L}$ - сильн. вып.

$\varphi(\lambda)$ (сопряж. по Фенхелю–Лежандру) $\searrow$ $\frac{1}{\mu}$ - липш. град.

$\parallel$

$\max\limits_{x \in \mathbb{R}^d} \{ \langle \lambda, x \rangle - f(x) \}$

$\min\limits_{\substack{Ax = b \\ x \in Q}} f(x) \to$ μ-сильн. вып. в 2-норме.

$$\varphi(\lambda) = \max_{x \in Q} \{ \langle \lambda, b - Ax \rangle - f(x) \}$$

$$\Downarrow\ L_\varphi = \frac{\lambda_{max}(A^T A)}{\mu} - \text{липш. град.}$$

Даже если $f(x)$ L_f-липш. град имеет, то $\varphi(\lambda)$ - не обяз. сильн. вып.

$$\begin{cases} f(x(\lambda)) - f(x_*) \leq \langle \lambda, \nabla\varphi(\lambda) \rangle = \varepsilon_f \\ \|Ax(\lambda) - b\|_2 \leq \|\nabla\varphi(\lambda)\|_2 \end{cases}$$

$$\frac{1}{2L_\varphi} \|\nabla\varphi(\lambda)\|_2^2 \leq \varphi(\lambda) - \varphi(\lambda_*) = \varepsilon_\varphi$$

$$\|\nabla\varphi(\lambda)\|_2 \sim \sqrt{L_\varphi \varepsilon_\varphi}$$

$$\sqrt{L_\varphi \varepsilon_\varphi}\cdot R_\lambda = \varepsilon_f$$

$$\Downarrow$$

$$\varepsilon_\varphi \sim \varepsilon_f^{\,2}$$

Как решить проблему?

Идея $\to$ решать двойств. задачу за линейное время, т.е. $N(\varepsilon_\varphi) \sim \ln \varepsilon_\varphi^{-1}$.

методы отсечений (типа метода эллипс.)

использ. сильн. вып. φ (если есть если нет регуляр. (в гладк. случае)

Регуляризация двойств. задачи

$$\varphi^\mu(\lambda) = \varphi(\lambda) + \frac{\mu}{2}\|\lambda\|_2^2 \to \min_{\lambda\in\mathbb{R}^m}$$

$$\underbrace{\langle\lambda,\nabla\varphi(\lambda)\rangle}_{\varepsilon} \le \frac{L_{\varphi^\mu}}{\mu}\overbrace{\Big(\varphi^\mu(\lambda) - \underbrace{\min_{\lambda\in\mathbb{R}^m}\varphi^\mu(\lambda)}_{\varphi^\mu(\lambda_*^\mu)}\Big)}^{\varepsilon_\varphi}$$

$$L_\mu = L_\varphi + \mu$$

$$\|\lambda_*^\mu\|_2^2 \le \frac{2}{\mu}\left(\varphi(0) - \varphi(\lambda_*)\right)$$

Д-во. $\varphi^\mu(\lambda) - \varphi^\mu(\lambda_*^\mu) \ge \frac{1}{2L_{\varphi^\mu}}\|\nabla\varphi^\mu(\lambda)\|_2^2 =$

$$\boxed{\begin{aligned}&f(x) \to \frac{\text{Напомним:}}{L\text{-гладк.}}\\ &f(x) - f(x_*) \geq \\ &\geq \frac{1}{2L_f}\|\nabla f(x)\|_2^2\end{aligned}}$$

$$= \frac{\|\nabla \varphi(\lambda) + \mu\lambda\|_2^2}{2L_{\varphi^\mu}} \geq$$

$$\geq \frac{\mu \langle \nabla \varphi(\lambda), \lambda\rangle}{L_{\varphi^\mu}}$$

$$\frac{\mu}{2}\|\lambda_*^\mu\|_2^2 = \frac{\mu}{2}\|\lambda_*^\mu - 0\|_2^2 \leq \varphi^\mu(0) - \varphi^\mu(\lambda_*^\mu) \leq$$

$$\frac{\mu}{2}\|0 - \lambda_*^\mu\|_2^2$$

$$\leq \varphi(0) - \varphi(\lambda_*)$$

$$\varphi^\mu(\lambda_*^\mu) = \varphi(\lambda_*^\mu) + \frac{\mu}{2}\|\lambda_*^\mu\|_2^2 \geq \varphi(\lambda_*) + \underbrace{\frac{\mu}{2}\|\lambda_*^\mu\|_2^2}_{\geq 0} \geq$$

$$\geq \varphi(\lambda_*)$$

<u>Замечание</u> 1) $\varphi(0) - \varphi(\lambda_*) \leq \min\limits_{\substack{Ax=b\\ x\in Q}} f(x) - \min\limits_{x\in Q} f(x)$

2) $\|\lambda_*^\mu\|_2 \leq \|\lambda_*\|_2$.

$\mu \sim \varepsilon$ – желаемой точности

$\varepsilon_\varphi \sim \varepsilon \cdot \mu$

это следует из того, что $\|\nabla\varphi(\lambda)\|_2 \leq \|\nabla\varphi^\mu(\lambda)\|_2 +$
$+ \mu\|\lambda\|_2$

$$\|Df(\lambda)\|_2 \sim \varepsilon$$

$$\Downarrow$$

$$\|Df^{\mu}(\lambda)\|_2 \sim \varepsilon/2$$

$$\mu\|\lambda\|_2 \sim \varepsilon/2 \implies \mu \sim \varepsilon/(2R)$$